부호화 효율을 고려한 고속 인터/인트라 모드 결정 알고리즘

정회원 김 지 웅*, 김 용 관*

Fast Inter/Intra Mode Decision Algorithm in H.264/AVC Considering Coding Efficiency

Ji Woong Kim*, Yong Kwan Kim* ***Regular Members***

요 약

H.264/AVC 비디오 부호화 표준은 부호화 효율을 향상시키기 위하여 기존의 부호화 표준들과는 다른 새로운 부호화 기법들을 사용한다. 그러나 새로이 채택된 여러 기법들로 인해 H.264/AVC 표준 부호기의 복잡도는 극단적으로 증가하게 되었다. 특히 율-왜곡 최적화 기법에 의한 H.264/AVC의 인터/인트라 모드 결정 방법은 부호기의 복잡도를 증가시키는 가장 큰 원인 중 하나이다. 본 논문에서는 매크로블록 모드 결정 과정의 복잡도 감소 및 전체 부호화 효율 측면에서 우수함을 나타내는 데 주안점을 두며, 이에 대한 고속 매크로블록 모드 결정 알고리즘을 제안한다. 제안한 알고리즘을 적용한 실험 결과, 고속 움직임 추정/보상과 결합한 경우 최대 80%의 부호화 시간을 감소시켰으며, 평균 8~10%의 비트율을 감소시키는데 반해 PSNR의 변화량은 미미하여 전체 부호화 효율 또한 상승시킬 수 있었다.

Key Words : H.264/AVC, Video coding, Rate-distortion optimization, Macroblock mode

ABSTRACT

For the improvement of coding efficiency, the H.264/AVC video coding standard employs new coding tools compared with existing coding standards. However, due to these new coding tools, the complexity of H.264/AVC encoder is greatly increased. Specially, Inter/Intra mode decision method of H.264/AVC using RDO(rate-distortion optimization) technique is one of the most complex parts in H.264/AVC. In this paper, we focus on the complexity reduction in macroblock mode decision considering coding efficiency. From the simulation results, the proposed algorithm reduce the encoding time by maximum 80% of total, and reduce the bitrate of the overall sequences by 8~10% on the average compared with existing coding methods.

Ⅰ. 서 론

2003년 최종 승인절차를 마친 H.264/AVC(advanced video coding) 동영상 부호화 표준은 이미 표준화 작업 초기에 기존 H.263과의 호환성을 고려하지 않고, 최대의 압축 효율을 최우선 목표로 하는 개발 방침을 내세웠다. 그 결과 H.264/AVC는 MPEG-2에 비해 약 40~50%, MPEG-4 ASP(advanced simple profile)에 비해 약 30~40% 성능이 향상된 것으로 보고되어 있다[1][2][3]. 특히 최근에 제안된 율-왜곡 최적화(RDO : rate distortion optimization) 방식은 H.26L의 압축 효율을 크게 향상시켰다[3][4].

반면, H.264/AVC에서는 복잡도의 증가를 고려하지 않고 압축 효율 향상을 최우선 목표로 삼았기 때문에 부호기 및 복호기의 계산량은 극단적으로 증가하여 높은 복잡도를 지니게 되었다[3]. 특히 새로이 추가된 부호화 기법들 중 다중 참조 프레임의 사용 및 1/4화소 단위의 움직임 추정과 가변 블록

* 호서대학교 정보통신공학과 영상통신 연구실 (ykkim@office.hoseo.ac.kr)
논문번호 : KICS2006-12-531, 접수일자 : 2006년 12월 14일, 최종논문접수일자 : 2007년 6월 22일

크기의 다중 예측 모드에 대한 반복적인 율-왜곡 비용(RDO_cost) 함수 계산은 부호기 및 복호기의 복잡도를 증가시키는 주된 요인이 된다. H.264/AVC 부호기에서는 최적의 인터/인트라 예측 모드를 선택하기 위해 다양한 인터/인트라 모드의 조합에 대해 율-왜곡 비용을 계산해야 한다. 율-왜곡 비용 함수에 사용되는 발생 비트량을 계산하기 위해서는 각 모드에 대한 DCT/Hadamard 변환 및 양자화/역양자화, 역 DCT/Hadamard 변환과정, 그리고 엔트로피 부호화 과정을 거치게 되므로 부호기의 복잡도를 크게 증가시키게 된다[4]. 따라서 전체 부호기의 복잡도를 감소시키기 위해서는 효율적인 고속 인터/인트라 모드 결정 알고리즘의 개발이 필수적이다.

최근에 인트라 모드 결정의 고속화를 위하여 소벨 마스크(Sobel mask)와 에지 히스토그램(edge histogram)을 이용하여 율-왜곡 계산을 필요로 하는 후보 모드의 수를 줄이는 알고리즘이 제안되었다[1]. [5]와 [7]에서는 특정 임계값을 이용하여 인트라 모드 결정 과정을 생략함으로써 인트라 모드 결정의 고속화를 이루었다. 또한 [6]에서는 다중 참조 프레임의 특성을 이용하여 참조 프레임의 결정을 고속화시키고 선택적 인트라 모드 결정법으로 전체 부호기의 계산량을 줄이는 알고리즘을 제안하였다. 그러나 이러한 기법들은 부호화 성능을 유지하면서 계산량을 충분히 감소시키지 못하는 단점이 있었다.

본 논문에서는 주변 블록의 모드와 현재 블록의 화소 분포를 이용하여 Intra4x4 모드 결정의 계산량을 최소화시킨다. 또한 Inter8x8 모드 내 최적의 서브 매크로블록(sub-macroblock) 모드 분포를 이용하여 Intra4x4와 Intra16x16 모드 결정 과정을 생략함으로써 인트라 모드 결정의 계산량을 감소시키는 알고리즘을 제안한다. 그리고 움직임 벡터와 참조프레임을 결정하기 위한 율-왜곡 비용 함수에서 발생될 수 있는 오차를 보상함으로써 전체 부호화 효율을 높이고, 부호기의 복잡도를 감소시킨다. 2장에서는 H.264/AVC 인터/인트라 예측 부호화 기법을 기술하고, 3장에서는 본 논문에서 제안하는 인터/인트라 모드 결정 알고리즘을 기술한다. 4장에서는 다양한 샘플 영상에 대한 실험 결과를 바탕으로 제안하는 알고리즘의 성능을 살펴보고, 5장에서 결론을 맺는다.

Ⅱ. H.264/AVC의 인터/인트라 예측 부호화

H.264/AVC의 예측 부호화 기법은 화면 내 예측(intra prediction)과 화면 간 예측(inter prediction)으로 구성된다. 화면 내 예측은 입력 영상의 공간적 중복성을 제거하는 부호화 기법이며, 화면 간 예측은 움직임 추정 및 보상을 통해 입력 영상의 시간적 중복성을 제거하는 부호화 기법이다[2][4].

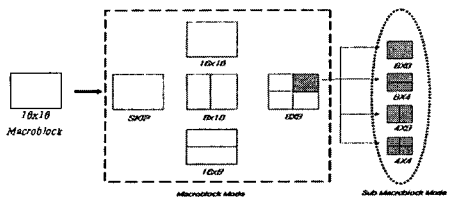
기존 부호화 표준에서는 화면 내 예측 부호화를 위해 변환 영역에서 DC 계수 또는 일부 AC 계수의 예측 부호화만을 수행하였다. 반면, H.264/AVC에서는 영상의 특성상 변환 영역의 계수보다는 공간 영역의 화소들 간 상관도가 높기 때문에 공간영역에서 화소 간 상관성을 최대한 제거하여 부호화 효율을 높이는 방법을 도입하였다[2]. 최적의 예측 모드를 결정하기 위해, 각 블록 단위별로 화면 내 복잡한 영역으로 구성된 Intra4x4 휘도 블록의 경우 9가지의 예측 모드 및 방향을 지원한다. 또한 비교적 평탄한 영역으로 구성된 Intra16x16 휘도 블록 및 Intra8x8 색차 블록에 대해서는 4가지의 예측 모드 및 방향을 지원한다. 이 때 예측을 수행하는 참조 블록은 복호(decoding)된 블록의 경계 화소를 이용하며, Intra4x4 휘도 블록의 경우, 총 8개의 방향성 모드와 DC 모드(방향성을 가지지 않는 평균치 예측 모드)로 구성된다.

화면 간 예측은 움직임 추정 및 보상을 이용하여 영상의 시간적 중복성을 제거함으로써 부호화 효율을 높이기 위한 부호화 기법이다. 움직임 추정 및 보상은 화면 간 예측에 있어 예측 부호화 후의 잔류 신호를 최소화하여 부호화 효율을 높이기 위해 반드시 필요하다. 움직임 추정 및 보상 과정에서 발생되는 움직임 정보는 각 화면 간 예측 모드별로 움직임 벡터(motion vector)와 참조 프레임(reference frame) 인덱스 정보, 그리고 해당 모드 번호로 구성된다[4]. H.264/AVC에서는 블록의 크기에 따라 Inter16x16, Inter16x8, Inter8x16, Inter8x8과 SKIP 모드의 5가지 화면 간 예측(inter prediction)모드를 지원하며, Inter8x8 모드의 경우 8x8 블록 안에서 8x8, 8x4, 4x8, 4x4 모드의 서브 매크로블록(sub-macroblock) 모드를 지원한다[2].

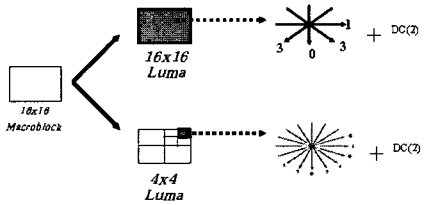
그림 1에서 볼 수 있듯이 H.264/AVC에서는 기존 동영상 부호화 표준에 비하여 다양한 예측 모드를 가지고 있으며, 각 모드별로 식 (1)의 율-왜곡 비용 함수가 최소값을 가지는 모드를 현재 매크로블록의 최적 모드로 결정한다[2][8].

$$J_{mode}(s,r,M \| \lambda_{mode}) = SSD(s,r,M) + \lambda_{mode} \times R(s,r,M) \quad (1)$$

s, r은 각각 원 매크로블록과 복호된 매크로블록



(a)



(b)

그림 1. H.264/AVC의 인터/인트라 예측 모드
(a) 인터 예측 모드 (b) 인트라 예측 모드 및 방향
Figure 1. Inter/Intra prediction mode of H.264/AVC
(a) Inter prediction mode (b) Intra prediction mode and direction.

의 화소값이며, M은 매크로블록 모드, 예측 방향 모드 또는 서브 매크로블록 모드이다. $R(s,r,M)$은 M에 해당하는 모드로 부호화했을 때 실제 발생되는 비트수이고, $SSD(s,r,M)$는 원 영상과 복호된 영상과의 차이에 대한 제곱의 합으로 정의된다. 또한 λ_{mode}는 라그랑지안(Lagrangian) 계수로 율-왜곡 함수에서의 접점의 기울기이다[3].

Ⅲ. 제안하는 인터/인트라 모드 결정 알고리즘

3.1 Intra4x4 모드 내 복잡도 감소

H.264/AVC의 인트라 모드에서 Intra4x4 모드의 경우, 총 9개의 예측 방향에 따라 최적 모드를 결정하고 가장 세분화된 블록에 대한 예측 부호화를 수행하기 때문에 상대적으로 높은 복잡도를 가진다. 그러나 그림 2와 같이 일반적으로 영상에서 주변 배경과 같은 영상 신호의 변화가 작은 영역에서는 현재 블록과 주변 블록의 화소 정보가 유사하다. 또한 Intra4x4 인트라 모드에서 현재 블록의 최적 예측 방향이 주변 블록의 예측 방향과 같다면 현재 블록과 주변 블록의 화소 정보 또한 비슷할 확률이 높다.

블록 간 유사성이 높을 경우 현재 Intra4x4 블록의 최적 모드를 결정하기 위하여 9가지 후보 모드 각각에 대한 율-왜곡 비용을 계산하기보다는 이미 결정된 주변 모드 정보만을 이용하는 것이 복잡도 측면에서 유리하다. 본 논문에서 제안하는 방식은 우선 현재 블록의 화소 분포와 주변 블록과의 화소 분포의 차이에 따른 블록 간 유사성을 판별하는 전처리 과정을 통해 후보 모드의 수를 줄이는 것이다.

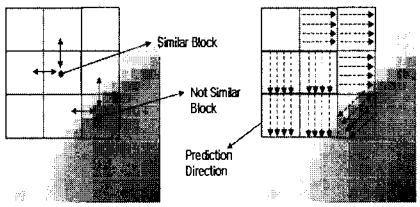


그림 2. 블록 간 유사성과 화소 정보에 따른 예측 방향 분포
Figure 2. Similar block and prediction direction distribution.

현재 블록 내 화소 분포도를 C_{dis}라 하고 현재 블록과 좌측 블록과의 화소 분포의 차이를 L_{cor}, 현재 블록과 상측 블록과의 화소 분포의 차이를 U_{cor}이라고 할 때 $C_{dis}, L_{cor}, U_{cor}$은 다음과 같이 계산된다.

$$C_{dis} = \sum_{x=0}^{T=15} |C_{avr} - CP_x| \tag{2}$$

$$L_{cor} = \sum_{x=0}^{T=15} |C_{avr} - LP_x| \tag{3}$$

$$U_{cor} = \sum_{x=0}^{T=15} |C_{avr} - UP_x| \tag{4}$$

이 때, C_{avr}은 현재 블록 내 화소 평균값이고, CP_x, LP_x, UP_x는 각각 현재, 좌측, 상측 블록의 x 인덱스에 해당하는 화소값이다. 그림 3은 블록 간 화소 분포를 계산하기 위한 현재 블록과 주변 블록의 화소 분포도를 나타낸다.

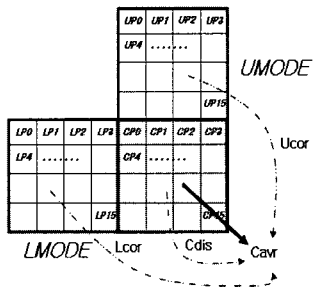


그림 3. 현재 블록과 주변 블록(좌측, 상측)의 화소 분포도
Figure 3. Pixel distribution of current block and neighboring blocks (left & up blocks).

$C_{dis}, L_{cor}, U_{cor}$을 계산한 후 현재 블록과 주변 블록과의 화소값 분포에 따른 블록 간 유사성을 판별하기 위한 L_{dif}과 U_{dif}은 식 (5)와 식 (6)에 의해 계산된다.

$$L_{dif} = |C_{dis} - L_{cor}| \tag{5}$$

$$U_{dif} = |C_{dis} - U_{cor}| \tag{6}$$

제안하는 Intra4x4 모드 내 복잡도 감소 알고리즘은 다음과 같다.

첫째, 만약 $L_{dif} < T$ AND $U_{dif} < T$ 인 경우,

즉, L_{dif}와 U_{dif}가 모두 임계값 T 보다 작을 경우 현재 블록의 화소 분포와 주변 블록과의 화소 분포는 상관도가 매우 높다고 판단되므로 현재 블록의 최적 모드는 이미 결정된 주변 블록의 모드로부터 예측 가능하다. 따라서 이 경우 Intra4x4 블록의 최적 모드를 결정하기 위한 선택 가능한 후보 모드는 그림 3의 LMODE와 UMODE로 압축된다. 또한 LMODE와 UMODE가 동일할 경우 선택 가능한 후보 모드의 수는 하나가 된다.

둘째, 만약 $L_{dif} > T$ OR $U_{dif} > T$ 인 경우,

현재 Intra4x4 블록의 최적 모드는 블록 내 화소값들의 방향성과 밀접한 관계를 가진다. [1]에서는 소벨 마스크(Sobel mask)와 에지 히스토그램(edge histogram)을 이용한 전처리 과정을 통해 후보 모드의 수를 줄임으로써 부호화 시간을 단축시킬 수 있었다. 그러나 소벨 마스크를 이용한 전처리 과정은 블록내의 정확한 방향성을 계산하기 위해 다소 복잡한 계산과정을 필요로 하며, 방향성을 검출하기 위한 과정에서 소수점 이하 연산을 필요로 하므로 부호화 시간 단축의 효과를 충분히 반영하지 못했다. 제안하는 방식은 그림 4와 같은 보다 간단한 구조의 4x4 정방형 마스크를 이용하여 Intra4x4 모드의 8가지 예측 방향 중 주요한 4가지 방향 정보를 구한다.

마스크 연산은 다음과 같이 계산되며, 최종적으로 선택 가능한 후보 모드는 연산 결과값이 최소가 되는 모드와 그림 4의 (b)에서 볼 수 있듯이 인접한 두 개의 모드 및 DC 모드로서 총 4개의 후보 모드를 가지게 된다.

$$Mode(0) = |a-m| + |b-n| + |c-o| + |d-p| \tag{7}$$

$$Mode(1) = |a-d| + |e-h| + |i-l| + |m-p| \tag{8}$$

$$Mode(3) = |c-i| + 2|d-m| + |h-n| \tag{9}$$

$$Mode(4) = |b-l| + 2|a-p| + |e-o| \tag{10}$$

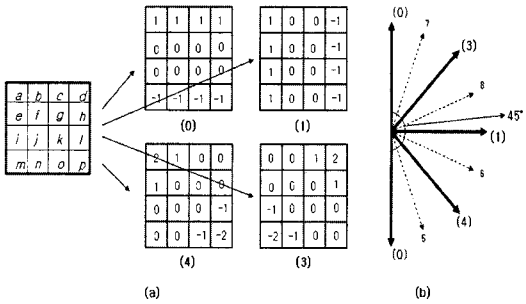


그림 4. (a) 4x4 정방형 마스크 구조 (b) 방향 정보
Figure 4. (a) 4x4 mask structure (b) direction information.

[1]에서 제안한 알고리즘의 경우, 선택 가능한 후보 모드의 수를 4개로 압축시킴으로써 율-왜곡 계산 횟수를 감소시켰다. 반면 제안하는 알고리즘의 경우 최소 하나의 후보 모드를 가지며 최대 4개의 후보 모드를 가지게 된다. 따라서 통상 4개 이하의 후보 모드를 가지게 되므로 율-왜곡 계산 횟수를 더욱 감소시킬 수 있다.

3.2 인트라 모드 결정 과정 생략

H.264/AVC에서 I-픽쳐(picture)를 제외한 P 또는 B 픽쳐의 경우 현재 매크로블록의 최적의 예측 모드가 인트라 모드로 결정될 확률은 3~4%정도로 미미하다. 따라서 H.264/AVC의 복잡도를 고려할 때 인트라 모드 검색을 꼭 필요한 경우에 한해 수행하도록 하는 것이 시간 측면에서 유리하다. [5]와 [7]에서는 특정 임계값에 기반 하여 인트라 모드 결정 과정을 생략하도록 함으로써 복잡도를 감소시켰다. 그러나 이 방식은 임계값의 변화에 민감한 단점을 가지고 있다. 따라서 본 논문에서는 특정 임계값을 이용하지 않고 Inter8x8 모드 내 서브 매크로블록(sub-macroblock) 모드 분포를 이용하여 선택적으로 인트라 모드 검색을 수행하는 방식을 제안한다.

H.264/AVC에서 영상의 배경과 같이 움직임이 작고 매크로블록 내 화소 분포가 비교적 균일한 경우 최적의 인트라 모드는 대부분 Intra16x16 모드로 결정된다. 또한 사물 영역과 같은 복잡하고 움직임이 많은 영역의 경우 최적의 인트라 모드는 Intra4x4 모드로 결정될 확률이 높다. 따라서 현재

매크로블록의 특성이 움직임이 작고 화소 분포가 비교적 균일하다고 판단될 경우, Intra4x4 모드에 대한 방향 결정 및 모드 결정 과정을 전부 수행하는 것은 복잡도 측면에서 비합리적이다. 제안하는 인트라 모드 결정 과정 생략 방식은 그림 5와 같은 Inter8x8 모드 내 서브 매크로블록의 최적 모드 분포를 이용하여 다음 단계에 따라 진행된다.

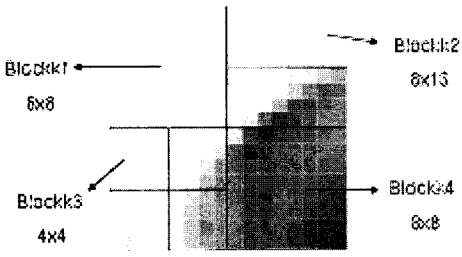


그림 5. 서브 매크로블록의 최적 모드 분포
Figure 5. Best mode distribution of sub-macroblock.

첫째, 만약 $Mode(Block1, Block2, Block3, Block4) = 8\times 8 Mode$

인 경우, 즉, Inter8x8 모드 내 서브 매크로블록의 최적 모드가 모두 8x8 블록 크기로 결정된 경우 현재 매크로블록은 8x8 블록 크기 이하의 세분화 과정을 필요로 하지 않는다. 왜냐하면 8x8 블록 크기 이하의 세분화 과정을 통해 왜곡 오차를 줄이더라도, 세분화된 블록 각각에 대한 비트 정보가 많아질 가능성이 높기 때문이다. 따라서 이 경우 인트라 모드 중 Intra4x4 모드에 대한 방향 결정 및 율-왜곡 계산을 생략한다.

둘째, 만약 $Mode(Block1, Block2, Block3, Block4) = 4\times 4 Mode$

인 경우, Inter8x8 모드 내 서브 매크로블록의 최적 모드가 모두 4x4 블록 크기로 결정된 경우 현재 매크로블록은 H.264/AVC의 가장 작은 블록 크기까지 세분화된다. 즉 현재 매크로블록은 영상의 사물영역과 같이 복잡하고 움직임이 많은 영역으로 판단할 수 있다. 따라서 이 경우 영상의 배경과 같은 비교적 평탄한 영역에 대한 화면 내 예측(intra prediction)을 수행하는 Intra16x16 모드에 대한 방향 결정 및 율-왜곡 계산 과정을 생략한다.

3.3 움직임 벡터와 참조 프레임 결정을 위한 율-왜곡 함수의 오차 보상

앞에서 기술한 바와 같이, H.264/AVC에서는 최적의 예측 모드를 결정하기 위하여 식 (1)의 비용 함수를 이용한다. 그러나 H.264/AVC의 인터 모드에서는 식 (1)의 비용 함수를 계산하기 전에 최적의 움직임 벡터와 참조 프레임을 먼저 결정해야 한다. H.264/AVC에서 최적의 움직임 벡터와 참조 프레임을 결정하기 위한 비용 함수는 식 (11)과 같이 식 (1)의 근사된 함수로 정의된다.

$$MVcost = D_{motion} + \sqrt{\lambda_{\mathrm{mode}}}\times R_{motion} \qquad (11)$$

H.264/AVC에서 최적의 움직임 벡터와 참조 프레임을 결정하기 위해 식 (11)과 같은 근사된 함수를 사용하는 이유는 H.264/AVC에서 채택한 움직임 추정/보상 특성에 기인한다. 실제 최적의 움직임 벡터와 참조 프레임을 찾기 위해서는 식 (1)의 비용함수를 이용하여 계산하는 것이 가장 정확하다. 하지만 탐색 범위 내에서 가능한 모든 영역에 대해 식 (1)의 율-왜곡 비용을 계산하는 것은 부호화 시간을 기하급수적으로 증가시키게 된다. 특히 H.264/AVC의 부호화 기법 중 1/4화소 정밀도 움직임 추정/보상 및 가변 블록 크기의 움직임 예측과 같은 특성을 고려할 때, 실제 식 (1)을 이용하여 율-왜곡 비용을 계산하는 것은 무리다. 따라서 H.264/AVC의 인터 모드에서는 최적의 움직임 벡터를 찾기 위하여 식 (1)을 이용하지 않고 보다 근사된 식 (11)을 이용하여 율-왜곡 측면에서 최소의 비용을 계산한다.

그러나 식 (11)은 실제 발생 비트율을 고려하지 않고 단지 움직임 벡터와 참조 프레임을 부호화하는데 필요한 비트수만을 고려하며, λ_{mode} 역시 근사된 값으로 제곱근을 이용한다. 따라서 식 (11)을 이용하여 결정된 움직임 벡터가 반드시 최종 부호화에서의 율-왜곡 측면에서 최적의 움직임 벡터라고 단정할 수 없으며, 만약 최적의 움직임 벡터가 아닐 경우, 결국 식 (1)을 이용한 최종 모드 결정에서 오차가 발생하게 된다. 본 논문에서는 이러한 오차를 실험적으로 검증하고 보상하기 위하여 식 (11)을 식 (12)와 같이 변경한 후, 다양한 영상에서 K의 변화에 따른 부호화 효율의 변화량을 살펴보았다.

$$MVcost = D_{motion} + K\sqrt{\lambda_{\mathrm{mode}}}\times R_{motion}, (K>1) \qquad (12)$$

식 (11)을 이용하여 결정된 움직임 벡터가 항시 최적이라면 K의 증가는 부호화 효율의 저하를 의미해야 한다.

표 1의 결과로부터 전체적으로 PSNR의 변화량은 미미한 수준에 그치는데 반해 K가 일정 범위 안에서 증가할수록 발생 비트율은 상대적으로 크게

표 1. K의 변화에 따른 전체 부호화 효율의 변화
Table 1. Coding efficiency as a function of K.

Sequence(QCIF)	Original		K=1.5		K=3		K=4		K=5		K=10	
	PSNR dB	Bit kbps	PSNR (dB)	Bit (%)	PSNR (dB)	Bit (%)	PSNR (dB)	Bit (%)	PSNR (dB)	Bit (%)	PSNR (dB)	Bit (%)
silent	35.753	84.698	0.008	-0.648	-0.047	-0.526	-0.072	-0.088	-0.109	0.5844	-0.193	5.355
foreman	35.726	138.95	-0.024	-1.604	-0.119	-2.099	-0.234	-1.873	-0.202	-0.8	-0.315	5.553
container	35.918	42.404	-0.013	-0.462	-0.056	0.004	-0.111	-0.203	-0.1	-0.212	-0.138	0.556
akiyo	38.553	23.187	-0.018	-0.294	-0.059	-1.212	-0.107	-0.763	-0.092	-0.815	-0.186	1.333
carphone	36.573	150.74	-0.042	-0.459	-0.145	-1.335	-0.272	-1.238	-0.259	-0.484	-0.378	3.949

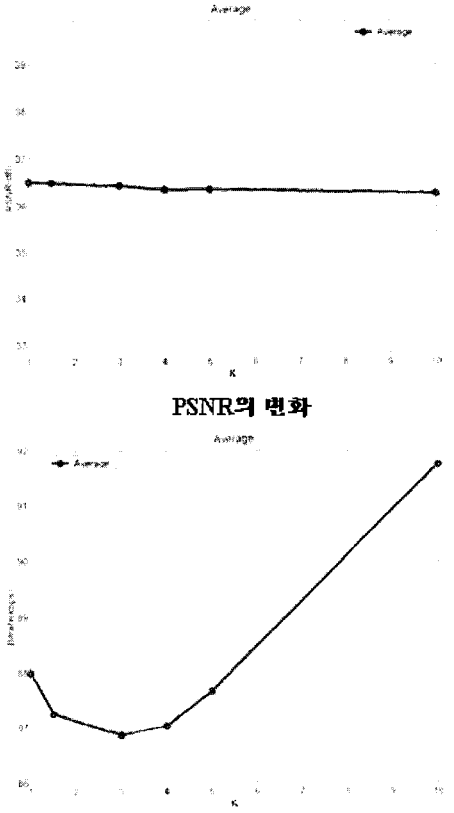


그림 6. K의 증가에 따른 PSNR과 비트율의 변화
Figure 6. PSNR and Bitrate as a function of K.

감소하는 것을 알 수 있다. 즉 그림 6에서 볼 수 있듯이 K의 증가에 따라 비트율이 상대적으로 큰 폭으로 감소함으로써 전체 부호화 효율은 오히려 증가한다.

전체 부호화 효율 측면에서 K의 증가는 λ_{mode}의 증가 또는 비트율 가중치 증가로 해석될 수 있다.

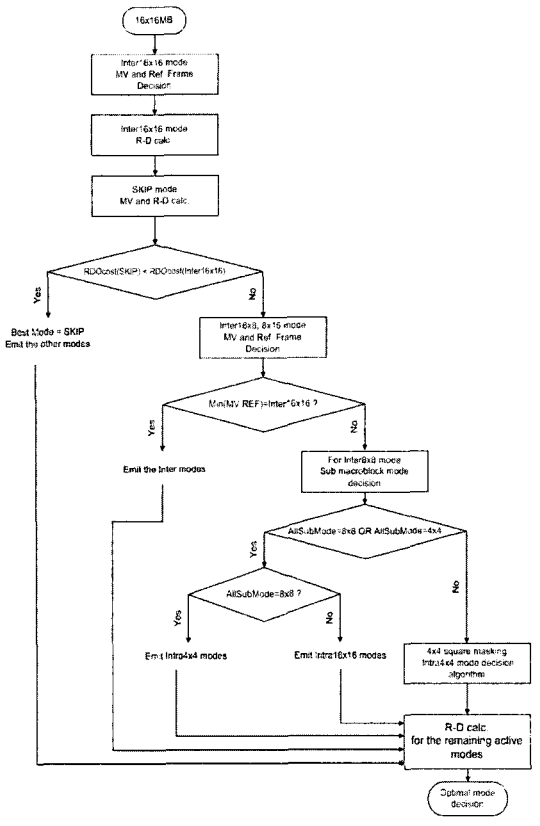


그림 7. 제안하는 알고리즘의 순서도
Figure 7. Flow chart of the proposed algorithm.

비트율 가중치 증가란 모드별로 보다 큰 블록에서 최적의 움직임 벡터가 결정될 확률을 높이는 것이다. 즉 상대적으로 낮은 발생 비트수를 가지는 모드에서 최적의 *MVcost*가 발생할 확률을 높이는 것이다.

H.264/AVC에서는 5개의 인터 모드와 2개의 인트라 모드를 제공하며, 이 중 상대적으로 낮은 복잡도와 발생 비트수를 가지는 모드는 SKIP, Inter16x16 모드이다. 두 모드는 특히 양자화기 값 Q가 클수록 발생 빈도수가 증가하게 되는데 이는 큰 양자화 값으로 인해 양자화 된 DCT 계수들이 0을 갖는 경우가 증가하여 발생 비트수가 감소하기 때문이다. 본 논문에서 제안하는 방식은 조기 SKIP 모드와 Inter16x16 모드에 대한 발생 빈도수를 증가시켜서 이하 세분화된 매크로블록에 대한 모드 결정 과정을 생략하도록 하는 것이다. 즉, 조기에 보다 세분화된 매크로블록 모드가 최적 모드로 결정될 경우를 배제시킴으로써 식 (12)의 K를 증가시키는 것과 같은 비트율 가중치 증가 효과를 나타낼 수 있다. 따라서 최종 모드 결정에서 근사화된 식 (11)로부터 발생될 수 있는 오차를 보상하여 전체

부호화 효율을 향상시킬 수 있다. 제안하는 방식은 다음 단계에 따라 조기 SKIP, Inter16x16 모드의 발생 빈도수를 증가시키게 된다. 제안하는 방식에서는 SKIP 모드의 발생 빈도수를 늘리기 위하여 최초 SKIP 모드에 대한 율-왜곡 비용을 계산한 후, Inter16x16 모드의 율-왜곡 비용과 비교한다. 비교 결과 SKIP 모드의 율-왜곡 비용이 작을 경우 이하 모든 모드 결정 과정을 생략하고 최적의 모드는 SKIP 모드로 결정한다. 반대로 SKIP 모드의 율-왜곡 비용이 클 경우에는 Inter16x16, Inter16x8, Inter8x16 모드에 대해 식 (11)을 이용한 최적의 *MVcost*를 결정한다. 최적의 *MVcost*가 Inter16x16일 경우 이하 모든 인터 모드 결정 과정을 생략한다. 그림 7에서 제안하는 인터/인트라 모드 결정 알고리즘의 전체 순서도를 도시하였다.

Ⅳ. 실험 및 결과

4.1 실험 환경

제안하는 알고리즘의 성능을 평가하기 위하여 본 논문에서는 H.264/AVC 참조 소프트웨어 JM9.5를 사용하였고, 테스트 방법은 JVT에서 권고하는 시뮬레이션 조건[9]을 따르도록 하였다. 또한 기존에 제안된 방식과 본 논문에서 제안하는 알고리즘을 비교함으로써 제안하는 알고리즘의 성능을 분석하였다. 실험 결과에 따른 성능 분석을 위해 사용되는 PSNR의 변화량($\triangle PSNR$), 비트율 변화량($\triangle Bit$) 및 전체 부호화 시간의 변화량($\triangle Time$)은 각각 다음의 식 (13), 식 (14), 식 (15)에 의해 결정된다.

$$\triangle PSNR = PSNR_{Algorithm} - PSNR_{original} \quad (13)$$

$$\triangle Bit = \frac{Bitrate_{Algorithm} - Bitrate_{original}}{Bitrate_{original}} \quad (14)$$

$$\triangle Time = \frac{Time_{Algorithm} - Time_{Original}}{Time_{Original}} \quad (15)$$

$\triangle PSNR, \triangle Bit, \triangle Time$의 값이 음수일 경우 감소량을 나타내며 양수일 경우 증가량을 나타낸다.

4.2 실험 결과 및 고찰

제안하는 알고리즘 중 비트율 가중치 증가를 적용하기 전의 방식(Ⅲ-1과 Ⅲ-2)과 [1]에서 제안한 고속 인트라 모드 결정 알고리즘(F. Pan's)과의 비교 테스트를 수행한 결과는 표 2와 같으며 구체적인 부호화 옵션은 다음과 같다.

- Search range = ±32
- Rate Distortion Optimization = Use
- # of reference frames = 1
- Symbol mode = CABAC
- GOP structure = IPPP... (I_frame_rate=100)
- QP = 28, 32, 36, 40.

테스트 영상으로는 QCIF(176x144) 크기의 영상 3개와 CIF(352x288) 크기의 영상 3개를 사용하였다.

표 2. F. Pan의 알고리즘[1]과 제안하는 인트라 모드 결정 알고리즘과의 실험 결과 비교
Table 2. Simulation results of F. Pan's algorithm[1] and the proposed intra mode decision algorithm.

Sequences (QCIFx3, CIFx3)		F. Pan's 알고리즘			제안하는 인트라 모드 결정 알고리즘								
					QP=28			QP=32			QP=40		
		ΔPSNR (dB)	ΔBit (%)	ΔTime (%)	ΔPSNR (dB)	ΔBit (%)	ΔTime (%)	ΔPSNR (dB)	ΔBit (%)	ΔTime (%)	ΔPSNR (dB)	ΔBit (%)	ΔTime (%)
QCIF	foreman	-0.077	1.536	-21.800	0.025	-5.602	-24.098	0.052	-3.307	-20.706	0.107	-0.574	-17.804
	container	-0.091	1.803	-20.784	-0.015	-3.094	-23.517	0.005	-1.727	-21.485	0.058	1.104	-18.055
	akiyo	-0.067	1.226	-23.109	0.027	-4.395	-21.216	0.006	-0.968	-19.615	0.091	1.336	-17.543
CIF	mobile	-0.018	0.451	-27.665	-0.018	-4.192	-31.008	0.002	-5.350	-27.851	0.021	-3.088	-21.926
	paris	-0.023	0.504	-26.901	0.023	-6.594	-27.607	0.013	-7.276	-24.879	0.030	-5.871	-22.618
	tempete	-0.029	0.812	-26.723	-0.016	-4.526	-27.553	-0.010	-4.279	-24.754	0.016	-2.382	-22.010

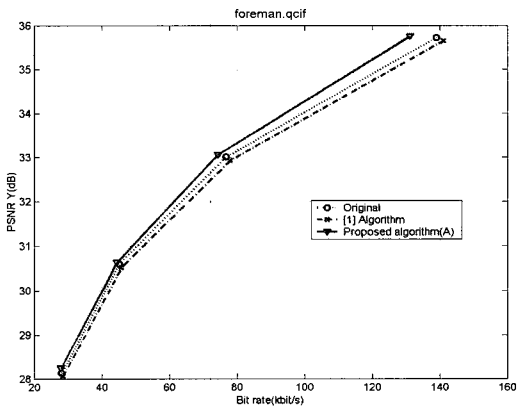


그림 8. foreman.qcif 영상에서 제안하는 인트라 모드 결정 알고리즘의 율-왜곡 곡선 성능 비교
Figure 8. RD-curve comparison of the proposed intra mode decision algorithm in foreman.qcif sequence.

실험 결과 전체 부호화 시간 성능에 있어서는 제안하는 인트라 모드 결정 알고리즘의 경우 F. Pan의 알고리즘[1]과 비교하여 비슷하거나 조금 높은 성능을 보였다. 부호화 효율 측면에서는 PSNR이 평균 0.03dB 정도의 미미한 변화를 가지는 반면, 발생 비

표 3. 비트율 가중치 적용 최종 알고리즘(C) 실험 결과
Table 3. Simulation results of the proposed final algorithm(C) with the increase of bitrate weighting effect.

Sequences	본 논문에서 제안하는 최종 Algorithm(C)											
	QP=28			QP=32			QP=36			QP=40		
	ΔPSNR (dB)	ΔBit (%)	ΔTime (%)	ΔPSNR (dB)	ΔBit (%)	ΔTime (%)	ΔPSNR (dB)	ΔBit (%)	ΔTime (%)	ΔPSNR (dB)	ΔBit (%)	ΔTime (%)
Foreman	-0.004	-6.2	-26.73	0.05	-6.21	-27.71	0.097	-6.56	-32.01	0.213	-11.85	-29.79
Container	0.015	-11.28	-36.76	0.005	-14.69	-35.66	0.071	-16.64	-35.09	0.068	-18.55	-34.54
News	0.052	-14.21	-34.18	0.054	-12.60	-34.36	0.157	-10.34	-33.28	0.106	-14.81	-33.62
Silent	-0.015	-7.6	-34.73	-0.029	-7.23	-32.31	-0.018	-8.08	-32.14	0.055	-10.48	-32.66
Carphone	-0.009	-7.84	-22.27	0.036	-8.22	-28.01	0.123	-7.45	-28.97	0.053	-10.9	-30.9
Salesman	-0.047	-12.92	-36.39	0.019	-12.02	-34.79	0.032	-11.91	-34.13	0.005	-14.48	-34.37
Claire	0.193	-12.37	-33.12	0.189	-12.71	-33.23	0.344	-13.21	-33.38	0.282	-16.64	-33.69

트율은 대부분의 영상에서 평균 3%이상 감소하였다. 이는 Inter8x8 모드 내 서브 매크로블록 모드 분포를 이용하여 인트라 모드 결정 과정을 생략함으로써 보다 적은 발생 비트수를 가지는 Intra16x16모드의 선택 빈도수를 증가시켰기 때문이다. 그림 8은 foreman.qcif 영상에서 제안하는 인트라 모드 결정 방식과 F. Pan등이 제안한 알고리즘과의 부호화 효율 성능에 따른 율-왜곡 (RD : rate distortion) 곡선을 나타내고 있다.

동일한 부호화 옵션을 사용한 경우, Ⅲ-3에서와 같은 비트율 가중치 증가 효과를 적용한 최종 알고리즘(C)의 성능을 테스트한 결과는 표 3과 같다. 표 3에서 알 수 있듯이, 본 논문에서 최종적으로 제안하는 기법은 전체 부호화 시간을 평균 28% 이상 단축시킬 수 있었다. 또한 부호화 효율 측면에서 PSNR은 평균 0.06dB의 미미한 변화량을 가지는 반면, 발생 비트율을 평균 8%이상 감소시키기 때문에 전체 부호화 효율을 크게 향상시킬 수 있었다. 특히 container, news와 같이 움직임이 작고 배경이 일정한 특성을 지니는 영상에서는 평균 10%이상 큰 폭으로 비트율을 감소시킬 수 있었다. 이러한 결과는 결국 Ⅲ-3 에서 기술한 바와 같이 식 (11)의 근사된 함수를 사용함으로써 발생되는 오차를 최대한 보상하여 잘못된 움직임 벡터로 인해 최적 모드가 잘못 결정될 경우를 최소화한 결과이다. 제안하는 최종 알고리즘과 [6]에서 제안하는 고속 모드 결정 방식과 비교 실험한 결과를 표 4에 나타내었으며, 이 때 사용된 부호화 옵션은 다음과 같다.

- # of reference frames = 5
- Fast motion estimation = Use
- Rate Distortion Optimization = Use
- Symbol mode = CAVLC
- GOP structure = IPPP...
- Search range = ±32
- QP = 28, 32, 36, 40.

표 4에서 볼 수 있듯이 전체 부호화 시간은 [6]에서 제안한 방식과 비교할 때 평균 15%이상의 부호화 시간을 단축시킬 수 있었다. 또한 부호화 효율 측면에서는 평균 0.04dB의 PSNR이 증가하였으며, 비트율의 경우 최대 16%까지 감소시킬 수 있었다.

제안하는 최종 알고리듬과 [6]에서 제안한 알고리즘의 부호화 효율 성능 비교를 위해 율-왜곡(RD) 곡선을 그림 9에 나타내었다.

그림 9에서 알 수 있듯이 제안하는 최종 인터/인트라 모드 결정 알고리즘은 식 (11)의 근사된 함수를 최대한 보상함으로써 전체 부호화 효율을 큰 폭으로 향상시킬 수 있었다. 특히 고속 움직임 예측과 결합하여 사용할 경우 부호화 시간을 약 80%까지 크게 단축시킬 수 있었다.

표 4. [6]에서 제안한 알고리즘과 제안하는 인터/인트라 모드 결정 알고리즘과의 실험 결과 비교
Table 4. Simulation results of [6]'s algorithm and the proposed Inter/Intra mode decision algorithm.

Sequences (QCIF×3, CIF×2)		[6]에서 제안한 방식			본 논문에서 제안하는 최종 알고리즘(C)								
					QP=28			QP=32			QP=40		
		ΔPSNR (dB)	ΔBit (%)	ΔTime (%)	ΔPSNR (dB)	ΔBit (%)	ΔTime (%)	ΔPSNR (dB)	ΔBit (%)	ΔTime (%)	ΔPSNR (dB)	ΔBit (%)	ΔTime (%)
QCIF	container	0.002	1.05	-50.09	0.046	-13.09	-77.77	-0.005	-16.1	-80.10	0.05	-15.57	-83.34
	carphone	-0.056	1.08	-44.22	0.047	-5.93	-51.30	0.09	-5.79	-56.51	0.163	-9.03	-69.95
	coastguard	-0.056	1.44	-46.09	0.04	-6.71	-49.2	0.062	-8.09	-57.32	0.145	-11.16	-71.81
CIF	foreman	-0.056	1.87	-43.69	0.079	-4.49	-52.25	0.125	-6.09	-60.34	0.302	-10.34	-73.10
	football	-0.165	1.25	-41.55	-0.025	-0.81	-40.79	-0.045	-1.47	-45.9	0.016	-2.89	-58.08
	crew	-0.075	0.61	-42.10	-0.032	-1.18	-40.15	-0.022	-2.94	-47.64	0.122	-6.72	-68.96

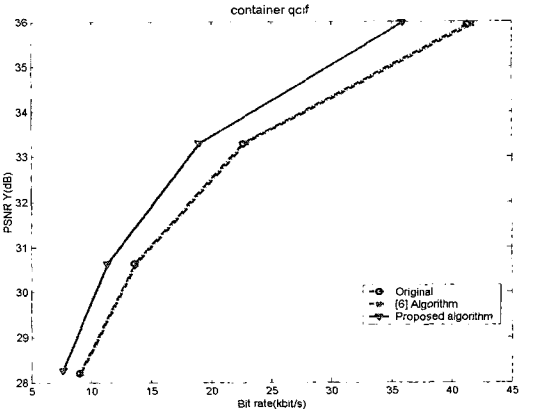


그림 9. container.qcif 영상에서 제안하는 최종 인터/인트라 모드 결정 알고리즘의 율-왜곡 곡선
Figure 9. RD-curve of the proposed final Inter/Intra mode decision algorithm in container.qcif sequence.

V. 결 론

기존의 율-왜곡 최적화 기법에 기반 하는 H.264/AVC의 모드 결정 방법은 비트율과 왜곡을 동시에 고려함으로써 부호화 효율을 높일 수 있었으나, 반복적인 율-왜곡 비용 계산에 따른 복잡도는 크게 증가되어 H.264/AVC 부호기에서 전체 부호화 시간을 증가시키는 주된 원인이 되었다. 따라서 본 논문에서는 이러한 부호기의 복잡도를 감소시키면서 동시에 부호화 효율을 향상시킬 수 있는 알고리즘을 제안하였다. 실험 결과로부터, 전체 부호화 시간에 있어 제안하는 알고리즘은 기존의 H.264/AVC 부호화 기법에 비해 IPPP 시퀀스에서 평균 30% 이상의 부호화 시간을 단축시켰으며, 비트율을 평균 8% 이상 크게 감소시킴으로써 전체 부호화 효율을 향상시킬 수 있었다. 앞으로 본 논문에서 제안한 고속 모드 결정 방법을 효과적으로 고속 움직임 예측 알고리즘과 결합할 경우, 고속 부호기 구현에 유용하게 이용될 수 있을 것으로 사료된다.

참 고 문 헌

[1] F. Pan, X. Lin, S. Rahardja, K. P. Lim, Z. G. Li, D. Wu, and S. Wu, "Fast Mode Decision Algorithm for Intraprediction in H.264/AVC Video Coding", *IEEE Trans. Circuits and Systems for Video Technology*, vol.15, no.7, pp.813-822, Jul. 2005.

[2] T. Wiegand, G. Sullivan, G. Bjontegaard, and A. Luthra, "Overview of the H.264/AVC video coding standard", *IEEE Trans. Circuits and Systems for Video Technology*, vol.13, no.7, pp. 560-576, Jul. 2003.

[3] J. W. Seok, B. H. Kim, and C. S. Cho, "The Research Trend of the H.264 Technology", *ETRI*, vol. 32, no. 1, pp. 25-34, 2006.2.

[4] ITU-T Rec. H.264/ISO/IEC 11496-10, "Advanced Video Coding", *Final Committee Draft, Document VT-E022*, September 2002.

[5] Y. Peng, C. T. Hyeyeon, M. T. Alexis, and B. Jill, "Fast Mode Decision and Motion Estimation for JVT/H.264", *IEEE international Conference on Image Processing*, vol.3, pp.853-856, September 2003.

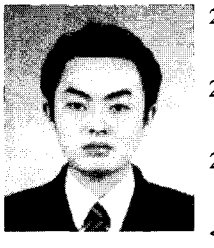
[6] L. Woongho, L. Jungho, C. Ikhwan, J. Dongseok, "Fast Mode Decision in H.264/AVC Using Adaptive Selection of Reference Frame and Selective Intra Mode", *Korea Institute Of Communication Sciences*, vol.31, no.3C, pp.271-278, 2006.3.

[7] K. Donghyung, J. Jechang, "A Fast MB Mode Selection Algorithm in the H.264 Standard", *Korea Institute Of Communication Sciences*, vol.30, no.1C, pp.61-72, 2005.1.

[8] G. Sullivan, T. Wiegand, and K. P. Lim, "Joint model reference encoding methods and decoding concealment methods", presented at the *9th JVT Meeting(JVT-I049d0)*, San Diego, CA, Sep. 2003.

[9] G. Sullivan,"Recommended simulation common conditions for H.26L coding efficiency experiments on low resolution progressive scan source material", presented at *the 14th VCEG-N81 Meeting*, Santa Barbara, CA, Sep. 2001.

김 지 웅(Ji Woong Kim) 정회원



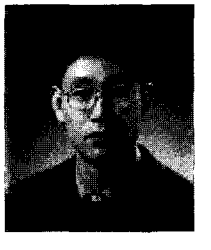
2005년 2월 : 호서대학교 정보통신공학과 학사

2007년 2월 : 호서대학교 정보통신공학과 석사

2007년 3월~현재:(주)우리기술연구원

<관심분야> 영상신호처리, ARM 기반 H/W 설계

김 용 관(Yong Kwan Kim) 정회원



1988년 2월 : 서울대학교 제어계측공학과 학사

1990년 2월 : 서울대학교 제어계측공학과 석사

1996년 8월 : 서울대학교 제어계측공학과 박사

1997년 3월~현재 : 호서대학교 정보통신공학과 부교수

2000년 3월~2001년 2월 : 캘리포니아 주립대(UCSB) 방문연구

<관심분야> 영상신호처리, 신호처리 SOC 설계, 멀티미디어통신